International Journal of ChemTech Research
CODEN (USA): IJCRGG, ISSN: 0974-4290, ISSN(Online):2455-9555
Vol.9, No.05 pp 887-895, 2016

Three dimensional QSAR analysis of Quinazolinone derivatives as EGFR Inhibitors

Jaheer Mohmed[1], Ranjith Reddy Palreddy[1], Raghav Rao Gundapaneni[1], Divya Jyothi Purohith[2], Shravan Kumar Gunda[2*]

[1]Department of Chemistry, Osmania University, Hyderabad, Telangana State, India
[2]Division of Bioinformatics, Osmania University, Hyderabad, Telangana State, India

Abstract : In the present investigation, a 3D QSAR studies had been employed for a series of quinazolinone derivatives which were acts as an EGFR inhibitors. CoMFA and CoMSIA, PLS fit methods were extensively used to predict the steric and electrostatic molecular field interactions for the better activity. The CoMFA and CoMSIA studies were achieved using a training set of 14 compounds. The 3D QSAR models of title compounds demonstrated that steric and hydrophobic interactions are dominant however, that substitution patterns are an important factor in determining activity. The actual assessed 3D-QSAR model has demonstrated a very good result, having r^2 value regarding 0.983 and also cross-validated coefficient q^2 value as 0.873. Molecular docking studies were useful in distinguishing a bioactive conformer and in addition a conceivable binding energies. The evaluation of CoMFA and CoMSIA contour maps furnished insight into the viable amendment of the molecules for higher activity.
Keywords: Protein tyrosine kinase (EGFR), 3D-QSAR, CoMFA, CoMSIA, 4-benzothiazole amino quinazoline dasatinib derivatives.

Introduction:

The protein kinase is the prevailing known family inside the human genome that involves more than 500 genes. Protein kinases are consequently ATP binding proteins that display profoundly rationed nature[1, 2]. Subsequently any functional deregulation of these enzymes results in illness states for example, cancer, diabetes, inflammation, cardiovascular disease, neurological issues and so on. EGFR belongs to human epidermal receptor family. It is a tans membrane glycoprotein. Glioblastoma multiforme, anal and lung cancer are marked by functional over activation of EGFR[1]. It plays very important role in signal transduction pathways. It also regulates key cellular functions such as adhesion, cell proliferation, migration, differentiation and survival[2]. The epidermal growth factor receptor (EGFR) is among the most well-known and demonstrated drug targets for cancer remedy.

Various numbers of small molecules i.e., EGFR kinase inhibitors had been evaluated in cancer scientific trials. A lot of compounds are still under assessment in scientific trials for the remedy of most cancers. Without further ado there are currently two primary classes of EGFR inhibitors that may be utilized as a part of tumor treatment and they're quinazoline derivatives and the pyrimidine derivatives. A lot of exploratory studies had been employed on 4-anilinoquinazoline derivatives which were powerful also, exceptionally specific inhibitors of epidermal growth issue (EGFR) phosphorylation on the ATP tying site

To find out new powerful EGFR inhibitors, scientist most likely must synthesize many compounds and screen their relating activities through cell based biological assay experiments which is generally time-ingesting and manpower expensive. In recent years there has been increment number of scientists are developing new EGFR inhibitors by employing various structure and ligand-based approaches like virtual screening[3], molecular docking[4], pharmacophore modeling[5] and QSAR[6, 7]. This methodology assists the scientist to foretell the activities of a series of newly designed drugs before making a choice regardless of whether to synthesize them and examine them.

Computational chemistry and computer aided drug design process have been extensively utilized now a day, not only lowering the price as well as in the proficiency with which they may be able to be designed. It is also additionally facilitates in discovery of key structural capabilities and assists within the evaluation of biological activity. In reality, structure-based, 3D-QSAR models will not be constrained to the ligand arrangement from only a solitary protein structure. In light of the ligand arrangement, we can make a sort of 3D-QSAR model from various protein structures.

In the present study, few QSAR models were created for anticipating inhibition interest of molecules towards EGFR. CoMFA and CoMSIA have a combination of reasonable molecular descriptors, statistical analysis and graphical representation of results. In the present investigation, we developed the 3D QSAR CoMFA[8] and CoMSIA[9] models on EGFR inhibitors in the expectation of getting a model that would account for the quantitative differences in biological activity seen in this series of EGFR derivatives and to capitalize upon the insights to design ligands with pronounced inhibitory potency and selectivity. These models are valuable to a constrained arrangement of molecules for a specific class like quinazoline-derivatives

Computational Studies:

Dataset:

In the present study, the biological activity data reported as IC_{50} for inhibition of EGFR[10] by the 4-benzothiazole amino quinazoline dasatinib derivatives as potential anti-tumor agents was used. Biological activities were reported in IC_{50} values were then converted into the corresponding pIC_{50} using the formula $pIC_{50} = -\log IC_{50}$.[11] The logarithmic transformation helps to get a symmetrically distributed data which is apt for the PLS regression analysis and it is important because log-dose response curve is linear about its middle region.

Molecular Modeling:

All molecular modeling studies were performed using the molecular modelling package SYBYL6.7[12] installed on a Silicon Graphics Work station. All the molecules were sketched and minimized by using Gasteiger-Huckel charges. Taken Tripos force field and 0.005 kcal/mol cutoff value was used. Later performed geometry optimization using MOPAC[13] interfaced with SYBYL. MMOK, ESP, NOMM was used during geometry optimization and MOPAC partial charges were computed.

Molecular Alignment:

Alignment plays very important role in 3D QSAR studies. One of the important assumptions wherein three dimensional quantitative structure activity relationship studies are based on that a geometric parallelism should present between the modeled structures of molecules and that of the bioactive conformation of the molecule. It indicates spatial alignment of compounds under study as one of the most delicate and determining factors in obtaining robustness and meaningful models[14]. In the present study the MOPAC (molecular orbital package) geometry optimized derivatives were aligned on the template by the routine ALIGN DATABASE command in SYBYL employing the divide and conquer strategy[15] as follows: molecules that have maximum common substructure for alignments. Finally these alignments were combined for subsequent CoMFA and CoMSIA study. Fig. (1) Shows the all aligned molecules.

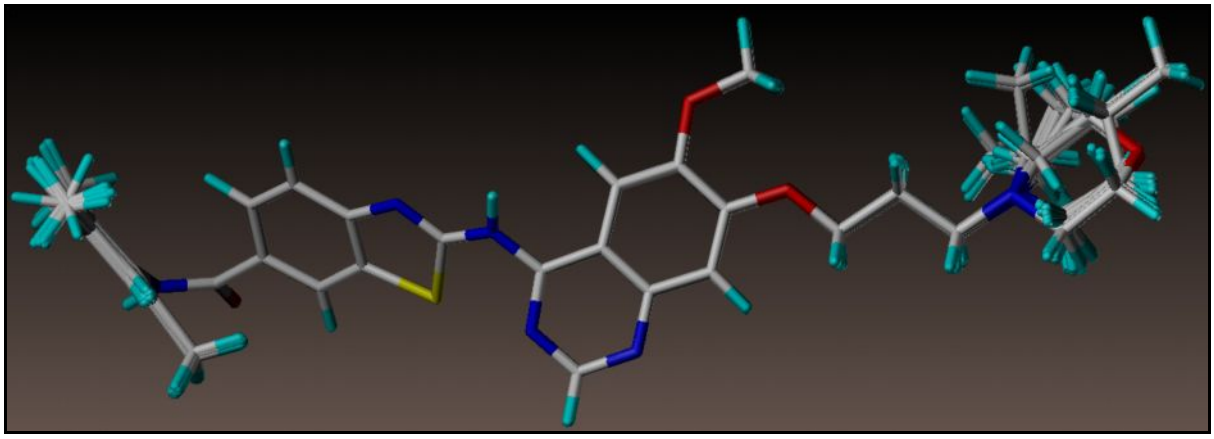


Figure 1. Possible binding conformations of 4-benzothiazole amino quinazoline dasatinib derivatives superimposed on the structural template (compound 11). The aligned molecules are shown in capped sticks.

CoMFA Interaction Energies:

The Comparative molecular field analysis (CoMFA) steric and electrostatic field parameters were calculated at each lattice intersection of a regularly spaced 2.0Å as grid dimensions. The van der Waals potential and columbic potential represents steric and electrostatic fields respectively. These two were calculated using the standard Tripo's force field. A distance dependent dielectric constant of 1.00 was used. sp^3 hybridized carbon atom with +1 charge served as probe atom to calculate steric and electrostatic fields. +30.0 kcal/mol steric and electrostatic contributions were truncated.[16] The cross-validation (CV) analysis was performed using leave-one-out (LOO) method. q^2 that resulted in optimum number of components and lowest standard error of estimate was considered. Equal weights for CoMFA were assigned to steric and electrostatic fields using CoMFA-standard scaling option.[17] To speed up the analysis a minimum column filtering value of 2.00 kcal/mol was used for the cross-validation. Further, final validation analysis was performed to calculate non cross-validated r^2 using the optimum number of components obtained from the leave one out cross validation analysis. Bootstrapping analysis for 100 runs was performed to assess the robustness and statistical confidence of the obtained models.

Table 1: Statistical analysis of 4-benzothiazole amino quinazoline dasatinib derivatives as EGFR inhibitors

	CoMFA		**CoMSIA**	
q^2	0.870		0.866	
r^2	0.983		0.994	
SEE	0.046		0.028	
F value	89.944		204.885	
CV	0.886		0.868	
Bootstrap	Mean	Stddev	Mean	Stddev
	0.032	0.032	0.025	0.026
	0.992	0.010	0.996	0.005
Field Contribtion (%)				
Steric	50.5		18.2	
Electrostatic	49.5		22.9	
Hydrophobic	-		25.2	
Donor	-		0.00	
Acceptor	-		33.7	

CoMSIA:

In the present study, we analyze the nature of 4-benzothiazole amino quinazoline dasatinib derivatives using 3D-QSAR (Three-dimensional quantitative structure–activity relationship) analysis. (CoMSIA) Comparative molecular similarity indices analysis was used. In CoMSIA, changes in ligand affinities are directly related to changes in molecular properties.[18] CoMSIA method is good at describing the intermolecular interactions (steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor) present at the molecular binding site. The method has been used to study the receptor-ligand interactions before and has proved to be of good predictivity.

Partial Least Square (PLS):

The CoMFA and CoMSIA analyses were performed using the partial least square (PLS) method. PLS regression technique is useful in common cases where the number of descriptors is comparable to or greater than the number of compounds and / or there exist other factors leading to correlations between variables. Biological activity is used as dependent variable and molecular descriptors as independent variable. The column filtering was set to 2.0 kcal/mol, to improve the signal-to-noise ratio. q^2 (conventional r^2) were performed by the Leave-One-Out (LOO) procedure, to determine the number of components (N). The cross-validated r^2 that resulted in number of components and lowest standard error of estimate were considered for further analysis. No-validation, cross-validation and finally bootstrapping analysis was performed to calculate conventional r^2 using the optimum number of components. Bootstrapping analysis for 100 runs was performed.

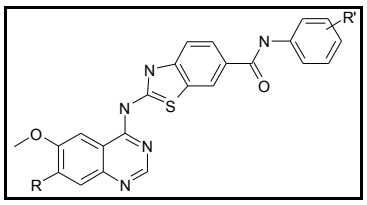
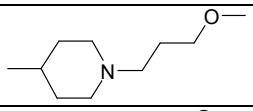
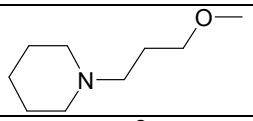
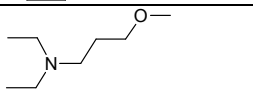
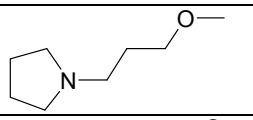
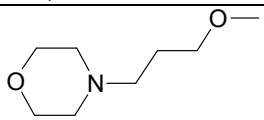
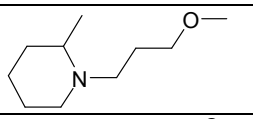
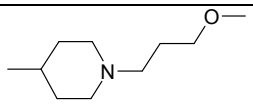
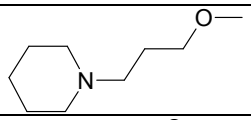
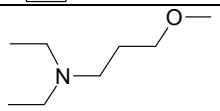
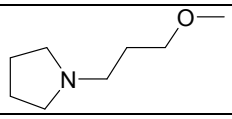
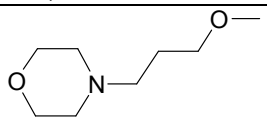
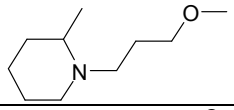
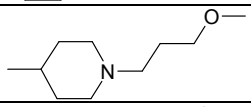
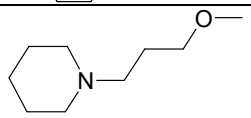
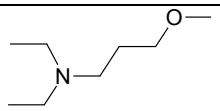
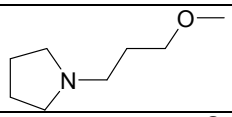
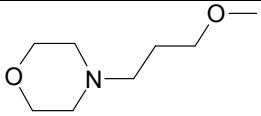
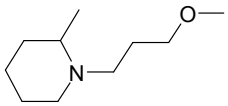


Table 2: Experimental, Predicted and Residual values of 4-benzothiazole amino quinazoline dasatinib derivatives used in Training set and Test set ; Bold and * indicates Test Set Compounds

C.No			pIC_{50}	CoMFA		CoMSIA	
				Predicted	Residual	Predicted	Residual
1*	2-Cl, 6-Me		**4.62**	**4.720**	**-0.100**	**4.650**	**-0.030**
2	2-Cl, 6-Me		4.54	4.677	-0.137	4.534	0.006
3*	2-Cl, 6-Me		**4.32**	**4.820**	**-0.500**	**4.650**	**-0.330**
4*	2-Cl, 6-Me		**4.80**	**4.430**	**0.370**	**4.380**	**0.420**
5	2-Cl, 6-Me		4.98	4.800	0.180	4.744	0.236
6	2-Cl, 6-Me		4.27	4.380	-0.110	4.330	-0.060
7	2,6-di-Me		4.87	4.778	0.092	4.818	0.052

8	2,6-di-Me		4.76	4.677	0.083	4.804	-0.044
9	2,6-di-Me		4.97	4.938	0.032	4.980	-0.010
10	2,6-di-Me		4.67	4.658	0.012	4.696	-0.026
11	2,6-di-Me		5.00	5.186	-0.186	5.233	-0.233
12	2,6-di-Me		4.59	4.509	0.081	4.547	0.043
13	2,4,6-tri-Me		4.44	4.480	-0.040	4.427	0.013
14	2,4,6-tri-Me		4.42	4.477	-0.057	4.359	0.061
15	2,4,6-tri-Me		4.59	4.561	0.029	4.517	0.073
16	2,4,6-tri-Me		4.31	4.322	-0.012	4.306	0.004
17*	2,4,6-tri-Me		**4.34**	**4.610**	**-0.270**	**4.470**	**-0.130**
18	2,4,6-tri-Me		4.18	4.172	0.008	4.175	0.005

Results and Discussion:

The 3D-QSAR CoMFA and CoMSIA studies were carried out using 4-benzothiazole amino quinazoline dasatinib derivatives, which are reported IC_{50} values on EGFR. 18 molecules were taken for the present study. All the 18 compounds were partitioned into a training set of 14 and a test set of 4 compounds as 1:3 ratio (1 percent in test set is and 3 percent in training set) were selected randomly. The ambiguity of ligand-receptor interactions in general, statistically robust models were obtained from the CoMFA and CoMSIA models. Experimental and predicted activities are given in table 2. Correlation between the predicted pIC_{50} values vs experimental Pic_{50} value for training and test set compounds are shown in Figure 2.

The CoMFA and CoMSIA PLS analysis is summarized in Table 1. The cross-validated correlation co-efficient is used as a measure of goodness of prediction. The non-cross-validated conventional correlation co-efficient indicates goodness of fit of a QSAR model. F value indicates for the degree of statistical confidence. A cross-validated correlation co-efficient of 0.870 was obtained using 5 as optimum number of components and 2.0 kcal/mol column filtering was used for the present model. The r^2cv obtained indicates a good internal predictive ability of the models. The models developed also exhibited a very good non-cross validated correlation co-efficient r^2 of 0.983. The Test set compounds are used to evaluate the external predictive capabilities of QSAR models. 4 compounds were selected in test set randomly were set-aside during model development. Further, a bootstrapping analysis was done for 100 runs. The r^2bs value obtained 0.992 of

bootstrapping by 100 runs which further supports the statistical validity of the developed models and absence of chance correlation. The contributions of steric to electrostatic fields were found to be 50.5% for steric and 49.5% for electrostatic. Both steric and electrostatic contributions are present in almost equal percentage.

The optimum CoMSIA model was derived with the combination of steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor field contribution using Gasteiger-Hückel charge with 2.0 Å grid space. Leave one out analysis gave the cross-validated q^2 of 0.866 with 6 components and column filtering was set to 1.0 kcal/mol. Non-cross-validated PLS analysis resulted in a correlation coefficient r^2 of 0.994, F= 204.885, with an standard error of estimate 0.028. Later we performed bootstrapping analyses to evaluate the robustness and statistical confidence of the final models (r^2 bootstrapping = 0.996, StdDev= 0.005). Statistical results obtained from the constructed model verified the predictive ability of the model. The predictive ability of the developed CoMSIA model was assessed by the test set (four molecules) predictions, which were excluded during model generation. Predicted and experimental activities and their residual values of all inhibitors are shown in Table 2.

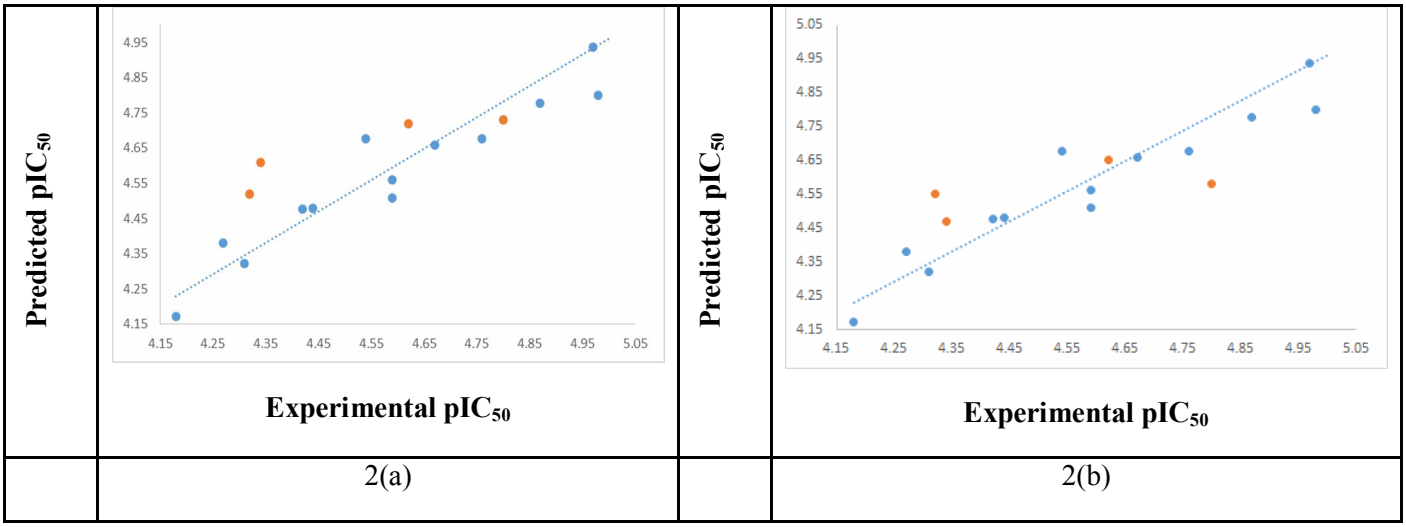


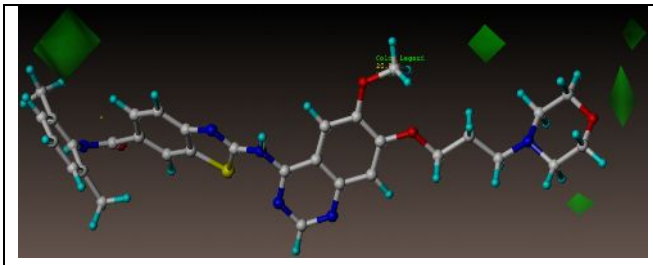
Fig. 2. Correlation between the predicted values vs experimental value for training and test set compounds by CoMFA 2(a), and CoMSIA 2(b).

Contour Analysis:

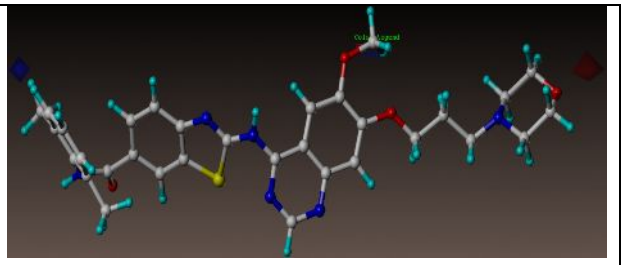
In SYBYL, steric interactions are displayed by green and yellow colored contours while electrostatic interactions are represented as red and blue contours. Green contours indicate where sterically bulkier groups are anticipated to increase the activity, whereas the yellow contours are used to decrease the points where bulkier groups could lower the biological activity. The electrostatic red contours show where the presence of a negative charge is expected to increase the biological activity whereas the blue contours indicate where inserting positive charge is expected to better the experimental activity.

CoMFA:

The steric contour of the CoMFA-EGFR model is displayed in Fig. 3 (a). From the Fig. 3 (b) it is clear that the plot shows only green contours extending over the R' position, increasing the steric crowding group at this position is expected to increase the inhibitory potency of the molecule. This fact is consistent with the better activities of the relatively bulkier compounds as compared to most active compound (11). The steric plot does not show any yellow plots in the present model. These contours are in conformity with the experimental observation. The plot also shows four small green contours extending over R position also.



3 (a)



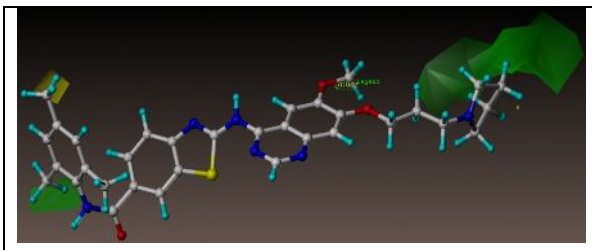
3(b)

Figure 3: CoMSIA (a) Steric (Green favorable, yellow unfavorable) (b) Electrostatic (Blue favorable, Red unfavorable contour maps onto compound 11.

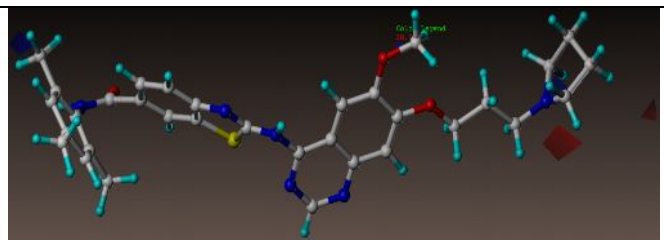
The steric and electrostatic contour map for CoMFA-EGFR model is displayed in Fig. 3 (b). It shows only one small red polyhedra at end of the R group. This is an indication where increase in negative charge is expected to improve activity of the molecule. Also observed a small blue polyhedron at R position. The most active molecule shows a positive interactions with the blue region here and hence accounting for the relatively higher inhibitory potencies.

CoMSIA:

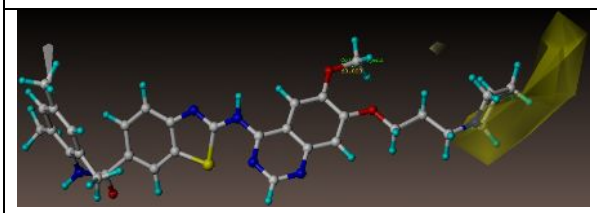
CoMSIA contour map was generated based on the atom-by atom matching (ligand-based) alignment method. CoMSIA contour maps shows regions where variations in the training set lead to increase or decrease in the activity. Steric interaction is represented by green and yellow contours, in which green colored regions indicate areas where increased steric bulk is associated with improved activity, and yellow regions suggest areas where increased steric bulk is unfavorable to activity. CoMSIA contour maps are displayed in Figure 4.



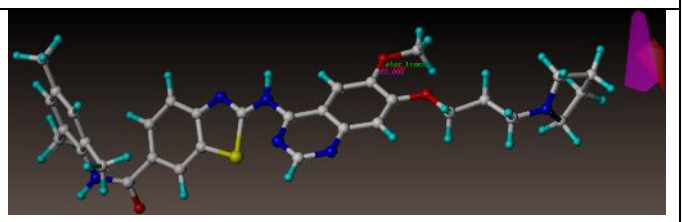
4 (a)



4 (b)



4(c)



4(d)

Figure 4: CoMSIA (a) Steric (Green favorable, yellow unfavorable) (b) Electrostatic (Blue favorable, Red unfavorable (c) hydrophobic (cyan favorable, yellow unfavorable) and (b) hydrogen bond acceptor (magenta favorable, red unfavorable) contour maps onto compound 11.

Figure 4(a) displays that there are two large green steric contours present at R and R' positions, which indicates that bulkier substituent is preferred at this position. The most active compound 11 with bulkier substituent at this position are more active because of this reason the compounds having bulkier substitution shows higher activity. Only one small yellow contour were observed near to R' position, the contour map

indicated that substitution of bulkier groups would decrease the activity. Lower activity compound having bulkier substitutions. Electrostatic interaction is represented by red and blue contours (Figure 4 (b)), among which blue colored contours show areas where more positively charged groups are favored, and red colored region indicate areas where groups with more negative charges are favored. Contour maps give us some general insight into the nature of the receptor-ligand binding region. A medium sized red colored contour situated close to the R position. The negative charges in these regions are very important for ligand binding, and electro negative group linked to this position will enhance the biological activity. The compound 11 having electro negative moiety shows more biological activity than other compounds. There is hydrogen bond acceptor contour maps are well correlated with the electrostatic contour map (Figure 4 (d)). Having a pink color contour at R position and a red color contour is also present at the same position. Hydrophobic contour map is well correlated with the existing steric contour map (Figure 4 (c)). According to the contour maps, steric, electrostatic, and H-bond acceptor requirements are playing major role in order to enhance the biological activities of the compounds. The contour map analyses indicated that bulky with more electronegative substitutions at the R' position is highly desirable to enhance the activity. Changes in these positions seems remarkably important to enhance the activity for EGFR inhibitors.

Conclusion:

Together with the statistical importance, good predictability and interpretation were achieved from the 3D-QSAR studies on a series 4-benzothiazole amino quinazoline dasatinib used in this present study. The built models provide valuable insight for identification of steric and electrostatic properties as well as hydrophobic, donor and acceptor hydrogen-bonding characteristics that could play important role for the future development of even more potent 4-benzothiazole amino quinazoline dasatinib based anti-tumor agents. Furthermore, molecular modeling methods were utilized to increase the understanding of substrate binding and catalytic mechanism. The ligand-based drug designing results are in coordination with each other and validate the present study for designing of more potential drug candidates of this class of compounds as antitumor agents.

References:

1. Pao W, Chmielecki J. Rational, biologically based treatment of EGFR mutant non-small-cell lung cancer. Nat Rev Cancer., 2010, 10, 760–774.
2. Singh H, Singh S, Singla D, Agarwal SM and Raghava GPS. QSAR based model for discriminating EGFR inhibitors and non-inhibitors using Random forest. Biology Direct, 2015, 10, 10.
3. La Motta C, Sartini S, Tuccinardi T, Nerini E, Da Settimo F, Martinelli A. Computational studies of epidermal growth factor receptor: docking reliability, three-dimensional quantitative structure-activity relationship analysis, and virtual screening studies. J Med Chem., 2009, 52, 964 -95.
4. Yadav IS, Nandekar PP, Srivastavaa S, Sangamwar A, Chaudhury A, Agarwal SM. Ensemble docking and molecular dynamics identify knoevenagel curcumin derivatives with potent anti-EGFR activity. Gene. 2014, 539, 82–90.
5. Gupta AK, Bhunia SS, Balaramnavar VM, Saxena AK. Pharmacophore modelling, molecular docking and virtual screening for EGFR (HER 1) tyrosine kinase inhibitors. SAR QSAR Environ Res., 2011, 22, 239–263.
6. Du H, Hu Z, Bazzoli A, Zhang Y. Prediction of inhibitory activity of epidermal growth factor receptor inhibitors using grid search-projection pursuit regression method. PLoS One., 2011, 6, e22367.
7. Chauhan JS, Dhanda SK, Singla D, Agarwal SM, Raghava GP. QSAR-based models for designing quinazoline/imidazothiazoles/pyrazolopyrimidines based inhibitors against wild and mutant EGFR. PLoS One. 2014; 9:e101079.
8. Cramer R.D, Patterson D.E, Bunce J.D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J. Am. Chem. Soc., 1988, 110, 5959–5967.
9. G. Klebe, Comparative Molecular Similarity Indices Analysis: CoMSIA, Perspect. Drug Discov. Des. 1998, 12(14), 87-104
10. Jin Cai, Min Sun, Xiaoqing Wu, Junqing Chen, Peng Wang, Xi Zong, Min Ji, Design and synthesis of novel 4-benzothiazole amino quinazolines Dasatinib derivatives as potential anti-tumor agents, Eur J Med Chem., 2013, 63, 702e712

11. Singh SK, Dessalew N, Bharatam PV. 3D-QSAR CoMFA study on indenopyrazole derivatives as cyclin dependent kinase 4 (CDK4) and cyclin dependent kinase 2 (CDK2) inhibitors. Eur J Med Chem., 2006, 41, 1310–1319.
12. Gunda SK, Kongaleti SF, Shaik M. Natural flavonoid derivatives as oral human epidermoid carcinoma cell inhibitors. Int J Comput Biol Drug Des., 2015, 8(1),19-
13. 39Hokuldsson A. PLS regression methods. J Chemo Metrics, 1988, 2, 211-228.
14. SYBYL6.7 is available from Tripos Associates Inc, 1699 S. Hanley Rd., St. Louis, MO 631444, USA.
15. Michael J.S. Dewar, Eve G.Z, Eamonn F.H, James J.P.S. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. J. Am. Chem. Soc., 1985, 107 (13), 3902–3909
16. Kim, K. H. In Molecular Similarity in Drug Design, Dean, P. M., Ed.; Blackie Academic and Profession, Glasgow, 1995, 303–305.
17. Ambrose Amin E, Welsh WJ. Three-dimensional quantitative structure-activity relationship (3D-QSAR) models for a novel class of piperazine-based stromelysin-1 (MMP-3) inhibitors: applying a "divide and conquer" strategy. J Med Chem. 2001, 44(23), 3849-3855.
18. Gunda SK, Narasimha SKM, and Shaik M. $P56^{lck}$ kinase inhibitor studies: a 3D QSAR approach towards designing new drugs from flavonoid derivatives, Int J Comput Biol Drug Des, 2014, 7 (2/3), 278-294

*****